

# Past quantum states

Søren Gammelmark, Brian Julsgaard, and Klaus Mølmer\*

*Department of Physics and Astronomy, Aarhus University,  
Ny Munkegade 120, DK-8000 Aarhus C, Denmark.*

(Dated: May 22, 2013)

A density matrix  $\rho(t)$  yields probabilistic information about the outcome of measurements on a quantum system. We introduce here the past quantum state, which, at time  $T$ , accounts for the state of a quantum system at earlier times  $t < T$ . The past quantum state  $\Xi(t)$  is composed of two objects,  $\rho(t)$  and  $E(t)$ , conditioned on the dynamics and the probing of the system until  $t$  and in the time interval  $[t, T]$ , respectively. The past quantum state is characterized by its ability to make better predictions for the unknown outcome of projective and weak value measurements at  $t$  than the conventional quantum state at that time. On the one hand, our formalism shows how smoothing procedures for estimation of past classical signals by a quantum probe [M. Tsang, Phys. Rev. Lett. **102**, 250403, (2009)] apply also to describe the past state of the quantum system itself. On the other hand, it generalizes theories of pre- and post-selected quantum states [Y. Aharonov and L. Vaidman, J. Phys. A: Math. Gen. **24**, 2315 (1991)] to systems subject to any quantum measurement scenario, any coherent evolution, and any Markovian dissipation processes.

PACS numbers: 03.65.-w, 03.65.Ta, 03.65.Ca

Quantum systems are described by wave functions or density matrices, which yield probabilistic predictions for the outcome of measurements performed on the systems. Following upon the rules laid out with the foundations of quantum theory in the 1920'es, the description of measurements on a so-called open quantum system has in the last few decades evolved into a well-established stochastic theory [1]. According to this theory, the density matrix  $\rho(t)$  evolves with time in a manner governed, on the one hand, by the system Hamiltonian and damping terms and, on the other hand, by the back action associated with the random outcome of measurements performed on the system or its environment. In this article, we introduce a new element in the quantum description of probed quantum systems: the past quantum state. While the density matrix  $\rho(t)$  yields predictions about the outcome of the measurement of any observable at time  $t$  conditioned on previous measurements, the past quantum state yields better predictions for the same measurement by being conditioned on all measurements carried out until the present time. The past quantum state is the state that we, based on what we know now, assign to a quantum system in the past. It is thus similar to the completely natural assignment of probabilities to past values of classical random quantities, e.g., for a Brownian particle detected at position  $x$  to have been at the position  $y$  at given earlier times. Here, we provide a generalization of the assignment of probabilities to past classical stochastic processes to the quantum case. Along with the definition and derivation of a past quantum state formalism, we shall answer the pressing questions: What does it mean to make predictions about the past? What are the new results and applications of a theory of past quantum states?

Consider an open quantum system subject to our continuous probing as illustrated in Fig. 1. We assume the

initial quantum state  $\rho_0$ , at time  $t = 0$ , and we probe the system until time  $T$  such that, conditioned on the measurement outcomes, the density matrix  $\rho(t)$  is given at any time  $t \in [0, T]$ . Specifically, if a different observer performs a measurement on the system at the intermediate time  $t$ , the density matrix  $\rho(t)$  provides the probability distribution of the possible measurement outcomes. Probing of the system after time  $t$  yields results that further refine our knowledge about the system at  $t$  and, indeed, there exists an *effect matrix*,  $E(t)$ , assuming the same Hilbert space dimension as  $\rho(t)$ , which depends on the dynamics and on the information acquired later than  $t$  until the present time  $T$  such that the pair of matrices,

$$\Xi(t) = (\rho(t), E(t)), \quad (1)$$

together enable better predictions than  $\rho(t)$  alone for the outcome of measurements carried out at time  $t$ . To discuss in a meaningful way what is meant by predicting a past measurement, we consider the setup shown in Fig. 1. Through an appropriately chosen interaction the physical property of interest is extracted at time  $t$  via coupling to another quantum system, an ancillary “meter”. This meter is stored “in a safe”, or it may be immediately measured and the result stored for later inspection. We show that  $\Xi(t)$  provides better predictions than  $\rho(t)$  of what will eventually be observed when the safe is opened. This qualifies  $\Xi(t)$  rather than  $\rho(t)$  to be associated to the *past quantum state* of the system.

The text book description of projective quantum measurements of a Hermitian operator is a specific case of general measurements associated with the action of different operators  $\hat{\Omega}_m$  that fulfill  $\sum_m \hat{\Omega}_m^\dagger \hat{\Omega}_m = \hat{I}$ , where  $\hat{I}$  is the identity and  $m$  is an index referring to the possible measurement outcomes [2]. For such a generalized measurement a suitable generalization of Born’s rule provides the probability at time  $t$  for observing the outcome  $m$ :

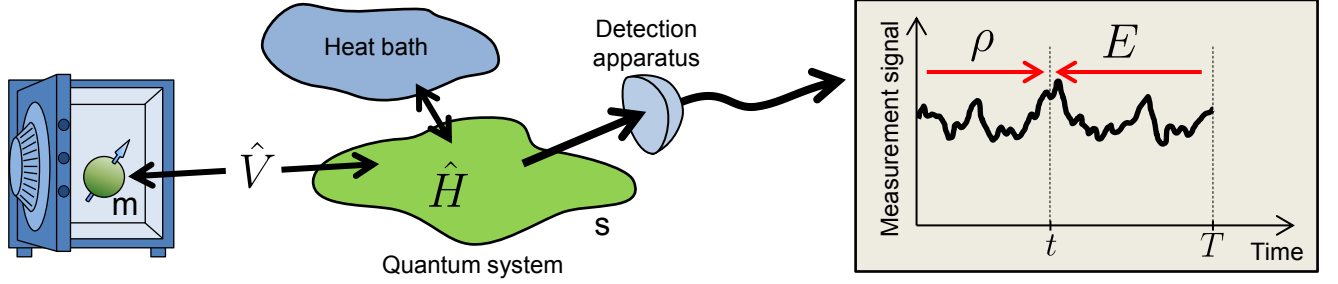


FIG. 1. Modeling a past measurement: A quantum system “s” is monitored during  $[0, T]$  and possibly subject to coherent evolution by the Hamiltonian  $\hat{H}$  and coupling to a heat bath. At time  $t$  a physical property of s is mapped to a meter “m” using an interaction  $\hat{V}$ . This meter, or the outcome of a measurement performed on the meter, is stored in a safe immediately after the interaction at time  $t$ . The matrices  $\rho$  and  $E$ , depending on the measurement signal before and after  $t$ , respectively, constitute the *past quantum state*, which predicts *better* the measurement outcome than  $\rho$  alone when the safe is opened at time  $T$ .

$p(m) = \text{Tr}(\hat{\Omega}_m \rho(t) \hat{\Omega}_m^\dagger) / \text{Tr}(\rho(t))$ . In the Supplementary Information [3] we prove that after further probing of the system until time  $T$ , the probability that the outcome is  $m$  depends on  $\Xi(t)$ , i.e., on both  $\rho(t)$  and  $E(t)$ :

$$p_p(m) = \frac{\text{Tr}(\hat{\Omega}_m \rho(t) \hat{\Omega}_m^\dagger E(t))}{\sum_m \text{Tr}(\hat{\Omega}_m \rho(t) \hat{\Omega}_m^\dagger E(t))}. \quad (2)$$

This formula is general and covers all possible measurement scenarios and any Markovian dynamical evolution, observed or non-observed, of our quantum system. As exemplified below and derived formally in the Supplementary Information [3],  $E(t)$  can be calculated backward in time following an *adjoint* equation very similar to the forward evolution of  $\rho(t)$ . In absence of probing,  $E$  retains its value  $E = \hat{I}$  for all times, and Eq. (2) reduces to the conventional expression since only our observations can further the knowledge of the state.

In the special case of a projective measurement of an observable  $\hat{A}$ , Eq. (2) applies with  $\hat{\Omega}_m = \hat{\Pi}_m$  denoting orthogonal projection operators on the eigenstates of  $\hat{A}$ . Past mean values, variances, and higher moments of  $\hat{A}$  then follow in the usual manner from  $p_p(m)$ . It is interesting to note that variances of past measurement outcomes will not necessarily obey Heisenberg’s uncertainty relations for non-commuting operators  $\hat{A}$  and  $\hat{B}$ . This, however, is not in violation of quantum mechanics since our probabilistic statements concern only the value of the observable actually measured, i.e., either  $\hat{A}$  or  $\hat{B}$ . In Ref. [4], a clever guessing game is suggested, where a person is given a spin-1/2 particle, and is free to measure either the  $x$ ,  $y$  or  $z$  component of the spin, and subsequently return the particle. By preparing the initial state and measuring the final state, it is possible to announce the outcome of any of these measurements without uncertainty. The past quantum state also permits a full analysis of such games.

Our general formalism permits also an analysis of so-called weak value measurements [5]. In this case the

strength of the interaction between a system observable  $\hat{A}$  and the meter can be parametrized by a small number  $\epsilon \ll 1$ , such that the disturbance resulting from the measurement is proportional to  $\epsilon^2$  and thus may be neglected. Nonetheless, averaged over sufficiently many experimental realizations the meter read-out will reveal the *mean value* of the system observable  $\hat{A}$  by the formula:  $\langle \hat{A} \rangle_w = \text{Tr}(\hat{A} \rho_p)$  where we have defined the *past density matrix*  $\rho_p = \rho E / \text{Tr}(\rho E)$  [3]. Here  $\rho$  and  $E$  are exactly the constituents of the past quantum state (1) and we thus provide a generalization of existing weak value expressions:  $\langle \hat{A} \rangle_w = \text{Re}\{\langle \varphi | \hat{A} | \psi \rangle / \langle \varphi | \psi \rangle\}$  or  $\langle \hat{A} \rangle_w = \text{Re}\{\text{Tr}(\hat{A} \rho_i E_m) / \text{Tr}(\rho_i E_m)\}$ , which apply, respectively, for a system initially prepared in a pure state  $|\psi\rangle$  or mixed state  $\rho_i$  and subsequently projected into the final state  $|\varphi\rangle$  or detected by a generalized measurement operator  $E_m$  [5, 6]. In these two examples  $\rho_p$  takes the form  $|\psi\rangle \langle \varphi| / \langle \varphi | \psi \rangle$  or  $\rho_i E_m / \text{Tr}(\rho_i E_m)$ , which has recently led to the recognition of these expressions as pre- and post-selected “connection states” [7]. In our general theory, the quantum state may at any time in the past be viewed as a pre-selection  $\rho$  by earlier measurement and a post-selection  $E$  by the later ones, with unitary evolution, dissipation processes, and direct measurement on the system also accounted for. We note that  $\rho_p$  is not described by a single evolution equation, but  $\rho$  and  $E$  must be calculated separately.

To illustrate the past state formalism and some of its results for a physical problem, let us turn to an example with a quantum system subject to coherent evolution, dissipation, and continuous homodyne-like monitoring. In such problems the usual quantum state  $\rho(t)$  is conditioned on the detection record until  $t \in [0, T]$  and it formally obeys the corresponding stochastic master equa-

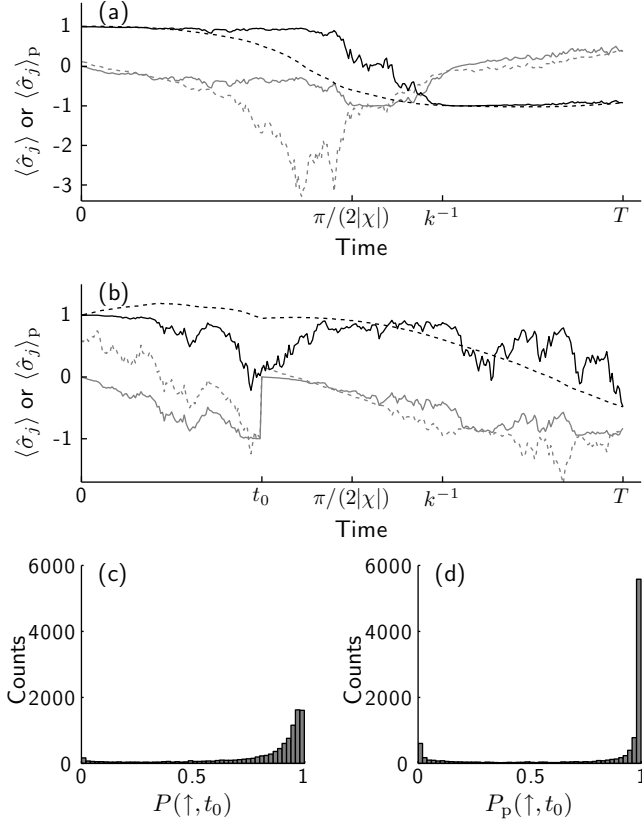


FIG. 2. Forward and past expectation values of a continuously monitored spin-1/2 particle: In panels (a) and (b) the curves show:  $\langle \hat{\sigma}_z \rangle$  [solid black],  $\langle \hat{\sigma}_x \rangle$  [solid gray],  $\text{Re}\{\langle \hat{\sigma}_z \rangle_p\}$  [dashed black], and  $\text{Re}\{\langle \hat{\sigma}_x \rangle_p\}$  [dashed gray]. In panel (b) an observer has performed a projective measurement of  $\hat{\sigma}_z$  on the system at time  $t_0$  without revealing the result. This leads to  $\langle \hat{\sigma}_x \rangle$  resetting to zero and the spectral boundary  $-1 \leq \text{Re}\{\langle \hat{\sigma}_z \rangle_p\} \leq 1$  automatically enforced. Panels (c) and (d) show the probability distribution for this projective measurement revealing spin-up based on  $\rho(t_0)$  and  $\Xi(t_0)$ , respectively, using 10,000 simulations.

tion [1, 8]:

$$d\rho_t = -i[\hat{H}, \rho_t]dt + \sqrt{\eta}(\hat{c}\rho_t + \rho_t\hat{c}^\dagger)dY_t + \sum_m (\hat{L}_m \rho_t \hat{L}_m^\dagger - \frac{1}{2}\{\hat{L}_m^\dagger \hat{L}_m, \rho_t\})dt, \quad (3)$$

where  $d\rho_t = \rho_{t+dt} - \rho_t$ ,  $\hat{H}$  is the interaction Hamiltonian,  $\hat{c}$  is the measurement observable,  $\hat{L}_m$  are Lindblad operators describing the dissipative coupling to the environment (including  $\hat{c}$  as one of the  $\hat{L}_m$ -terms),  $\eta$  is the detector efficiency, and  $dY_t$  is the measurement record. The effect matrix  $E$  solves a corresponding adjoint equation with final condition  $E = \hat{I}$  at  $T$  [3]:

$$dE_t = i[\hat{H}, E_t]dt + \sqrt{\eta}(\hat{c}^\dagger E_t + E_t \hat{c})dY_{t-dt} + \sum_m (\hat{L}_m^\dagger E_t \hat{L}_m - \frac{1}{2}\{\hat{L}_m^\dagger \hat{L}_m, E_t\})dt, \quad (4)$$

where  $dt$  is positive and  $dE_t \equiv E_{t-dt} - E_t$ , propagating backward from  $T$  to  $t$  using the same measurement record  $dY_t$  as in Eq. (3). We note that these equations are not trace preserving but can easily be adapted as such if required, e.g. for numerical evaluation.

For concreteness, we consider now a quantum two-level system subject to coherent driving with Rabi frequency  $\chi$  according to the Hamiltonian  $\hat{H} = \frac{1}{2}(\chi\hat{\sigma}_+ + \chi^*\hat{\sigma}_-)$  and to continuous probing of  $\hat{\sigma}_z$  through the measurement operator  $\hat{c} = \sqrt{k}\hat{\sigma}_z$ . Here  $\hat{\sigma}_j$  are Pauli spin operators and  $k$  is the measurement strength. Such measurement could be implemented by, e.g., polarization rotation of a radiation field coupled to the spin-1/2 particle [9]. We have performed simulations with  $\eta = 1$ , a pure initial state  $\rho_0 = |\uparrow\rangle\langle\uparrow|$ , and an imaginary  $\chi$  such that the coherent driving rotates the spin around the  $y$ -axis. In Fig. 2, the expectation value of the Pauli operators,  $\hat{\sigma}_z$  and  $\hat{\sigma}_x$ , for the quantum system are compared using the usual forward state  $\rho(t)$  and the past quantum state  $\Xi(t)$ .

The analysis in Fig. 2(a) applies to the case where the system has not been disturbed by further measurements by other observers. This implies that the past density matrix  $\rho_p(t)$  may be used to yield predictions for the outcome of weak value measurements of any system observable. We stress that despite its possible values beyond the interval  $[-1, 1]$ ,  $\langle \hat{\sigma}_j(t) \rangle_p = \text{Tr}(\hat{\sigma}_j \rho_p(t))$ , rather than  $\langle \hat{\sigma}_j(t) \rangle = \text{Tr}(\hat{\sigma}_j \rho(t))$ , represents the correct estimate of the disturbance of the meter system. For a spin-1/2 meter system, the real and imaginary parts of  $\langle \hat{\sigma}_j(t) \rangle_p$  correspond to mean rotation angles of the spin around different axes [10].

Fig. 2(b) exemplifies the case where an observer has performed a projective measurement of  $\hat{\sigma}_z$  at time  $t_0$  without revealing the result, and Eq. (2) enables a past prediction  $\langle \hat{\sigma}_z(t_0) \rangle_p$  of this outcome using  $\Xi(t_0)$ . For all other times  $t \in [0, T]$  the projective measurement must be taken into account in the evolution of  $\rho$  and  $E$ . To account for the decoherence by the measurement at  $t_0$ , we evolve the density matrix,  $\rho(t_{0+}) = \hat{P}_{\text{up}}\rho(t_{0-})\hat{P}_{\text{up}} + \hat{P}_{\text{down}}\rho(t_{0-})\hat{P}_{\text{down}}$  by the projection operators,  $\hat{P}_{\text{up}} = |\uparrow\rangle\langle\uparrow|$  and  $\hat{P}_{\text{down}} = |\downarrow\rangle\langle\downarrow|$ . Similarly, to obtain the value of the effect matrix  $E$  prior to  $t_0$ , we have to apply the operation  $E(t_{0-}) = \hat{P}_{\text{up}}E(t_{0+})\hat{P}_{\text{up}} + \hat{P}_{\text{down}}E(t_{0+})\hat{P}_{\text{down}}$ . It is particularly interesting to compare the predictions of the un-revealed measurement outcome using the conventional and the past quantum state formalism. For predictions associated with projective measurements,  $\langle \hat{\sigma}_z(t_0) \rangle_p$  is real and it remains within its spectral boundaries as it should to yield agreement with experiments. The result, however, differs from the prediction by the conventional density matrix, and we quantify this difference by the distribution of probabilities for the two-level spin direction to be registered as up or down. These distributions are shown for the conventional quantum state in Fig. 2(c) and for the past quantum state in Fig. 2(d). By assuming

that the most likely measurement result is the one occurring, we guess the outcome correctly with 88% probability by the conventional quantum state  $\rho(t_0)$ , while the past quantum state  $\Xi(t_0)$  yields the correct measurement outcome with 94% probability.

We note that if  $\rho(t)$  predicted a measurement outcome with certainty at time  $t$ , our later probing will never lead to disagreement with this prediction: If the system was in the  $m$ th eigen state  $|a_m\rangle$  of an observable  $\hat{A}$ , i.e.  $\hat{A}|a_m\rangle = a_m|a_m\rangle$ , at time  $t$ , then  $\rho(t) = |a_m\rangle\langle a_m| = \hat{\Pi}_m$  and the past probability for measuring the  $n$ th eigen state then becomes:  $p_p(n) = \delta_{n,m}$  according to Eq. (2) and the orthogonality relation  $\hat{\Pi}_m\hat{\Pi}_n = \hat{\Pi}_m\delta_{n,m}$ .

While the past quantum state enables a sharper prediction for the projective measurement shown in Fig. 2(c,d), we note that the weak value  $\langle\hat{\sigma}_z\rangle_p$  is actually smoother than the forward estimate  $\langle\hat{\sigma}_z\rangle$  in Fig. 2(a,b). In the Supplementary Information we show that while  $\langle\hat{\sigma}_z\rangle$  varies with noisy increments  $\propto dY_t \propto \sqrt{dt}$  due to Eq. (3), the changes in  $\langle\hat{\sigma}_z\rangle_p$  at time  $t$  are formally independent of  $dY_t$  and vary smoothly according to the integrated noise via  $\rho(t)$  and  $E(t)$ .

To exemplify an estimation process with mixed coherent and incoherent degrees of freedom, we consider in Fig. 3 a coherently driven, spontaneously decaying two-level atom which can jump incoherently between two sites,  $a$  and  $b$ . In Fig. 3(a) the position of the atom (as used in the simulations but in experimental realizations hidden from us) is shown along with the instants of detection of photons emitted from the two-level atom. Owing to different site environments, the coherent atom-dynamics is determined by different parameter values at each site. We can use the forward and past quantum states, based on the photo-detection record only, to estimate the location of the atom as shown in Fig. 3(b) and Fig. 3(c), respectively. As demonstrated in Refs. [11–14], a similar formalism enables the estimation of unknown classical perturbations, applied to a probed quantum system. These works generalize the so-called smoothing procedure applied in the classical probability theory of Hidden Markov Models (HMM) [15] to hybrid quantum classical systems. Such hybrid systems may, indeed, be embedded in a full quantum model, and we comment further on the formal similarities between our past quantum state and the smoothing procedure in classical HMM in the Supplementary [3]. The past quantum state estimate is both much less noisy and much more decisive, enabling a more accurate state estimation and, e.g., a better estimate of the  $a$ - $b$  jump rate. In addition to estimating such classical properties our formalism provides also a better estimation of the quantum system itself and we are currently investigating the application of our theory to the state of a photon field in a cavity [16] and to the state of a superconducting qubit [17–19].

The past quantum state  $\Xi(t)$  depends on events occurring in the future beyond the time  $t$ . While “spooky

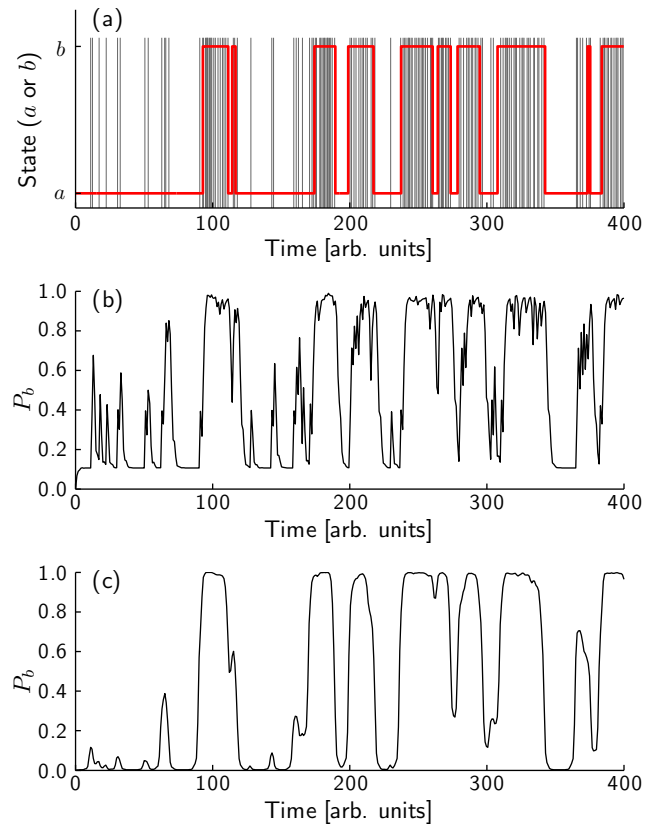


FIG. 3. Forward and past estimate of discrete position jumps of a driven two-level atom: In panel (a), the actual random position jumps of the atom are shown (red line) along with the photon detection times (vertical black). At position  $a$  (b), the atoms have a low (high) average emission rate. Panel (b) shows the probability for finding the atom at site  $b$  calculated using the forward density matrix, while panel (c) employs the past quantum state using the same data. Panel (c) is clearly more decisive and in much better agreement with the true state (red line in panel (a)).

action from the future” via post-selection has stimulated fascinating scientific debate [20], the predictions we make can be interpreted as correlations between system observables at the past time  $t$  and a probing signal acquired in the, also past, interval  $[0, T]$ . This distinction has also been made clear in the formal work on pre- and post-selected states [21, 22]. While the concept of a past quantum state is also central in these publications, our analysis aims less on the foundations of quantum theory and more on the general and explicit description of continuously monitored open quantum systems that are commonplace in laboratories today. Our theory applies to experiments on superconducting devices [17–19, 23]; semiconductor quantum dots [24]; NV-centers in diamond [25, 26]; nuclear spins in silicon [27]; trapped ions, atoms and molecules [28–30]; photons [16, 31, 32], and nanomechanical devices [33]. These systems, indeed, may be used for fundamental tests [19, 23, 25, 28, 31], but

they also hold the potential for application in precision probing [34] and quantum information science [2], and we hope our theory will stimulate further development of both technical and foundational aspects of the theory of open quantum systems.

This work was supported by the Villum Foundation.

---

\* moelmer@phys.au.dk

- [1] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, New York, 2010).
- [2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [3] “See supplementary material for details.”
- [4] L. Vaidman, Y. Aharonov, and D. Z. Albert, Phys. Rev. Lett. **58**, 1385 (1987).
- [5] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988).
- [6] L. M. Johansen and A. Luis, Phys. Rev. A **70**, 052115 (2004).
- [7] A. G. Kofman, S. K. Özdemir, and F. Nori, “Connection-state approach to pre- and post-selected measurements,” <http://arxiv.org/abs/1303.6031> (2013).
- [8] K. Jacobs and D. A. Steck, Contemp. Phys. **47**, 279 (2006).
- [9] Y. Takahashi, K. Honda, N. Tanaka, K. Toyoda, K. Ishikawa, and T. Yabuzaki, Phys. Rev. A **60**, 4974 (1999).
- [10] S. Wu and K. Mølmer, Phys. Lett. A **374**, 34 (2009).
- [11] M. Tsang, Phys. Rev. Lett. **102**, 250403 (2009).
- [12] M. Tsang, Phys. Rev. A **80**, 033840 (2009).
- [13] M. Tsang, Phys. Rev. A **81**, 013824 (2010).
- [14] M. Tsang, H. M. Wiseman, and C. M. Caves, Phys. Rev. Lett. **106**, 090401 (2011).
- [15] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (Cambridge University Press, New York, 2007).
- [16] S. Gleyzes, S. Kuhr, C. Guerlin, J. Bernu, S. Deléglise, U. B. Hoff, M. Brune, J.-M. Raimond, and S. Haroche, Nature **446**, 297 (2007).
- [17] M. Hatridge, S. Shankar, M. Mirrahimi, F. Schackert, K. Geerlings, T. Brecht, K. M. Sliwa, B. Abdo, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Science **339**, 178 (2013).
- [18] R. Vijay, D. H. Slichter, and I. Siddiqi, Phys. Rev. Lett. **106**, 110502 (2011).
- [19] J. P. Groen, D. Ristè, L. Tornberg, J. Cramer, P. C. de Groot, T. Picot, G. Johansson, and L. DiCarlo, “Partial-measurement back-action and non-classical weak values in a superconducting circuit,” <http://arxiv.org/abs/1302.5147> (2013).
- [20] Y. Aharonov, S. Popescu, and J. Tollaksen, Phys. Today **63**, 27 (2010).
- [21] Y. Aharonov and L. Vaidman, J. Phys. A: Math. Gen. **24**, 2315 (1991).
- [22] Y. Aharonov, S. Popescu, J. Tollaksen, and L. Vaidman, Phys. Rev. A **79**, 052110 (2009).
- [23] A. Palacios-Laloy, F. Mallet, F. Nguyen, P. Bertet, D. Vion, D. Esteve, and A. N. Korotkov, Nature Phys. **6**, 442 (2010).
- [24] I. T. Vink, K. C. Nowack, F. H. L. Koppens, J. Danon, Y. V. Nazarov, and L. M. K. Vandersypen, Nature Phys. **5**, 764 (2009).
- [25] G. Waldherr, P. Neumann, S. F. Huelga, F. Jelezko, and J. Wrachtrup, Phys. Rev. Lett. **107**, 090401 (2011).
- [26] J. Cai, F. Jelezko, M. B. Plenio, and A. Retzker, New J. Phys. **15**, 013020 (2013).
- [27] J. J. Pla, K. Y. Tan, J. P. Dehollain, W. H. Lim, J. J. L. Morton, F. A. Zwanenburg, D. N. Jamieson, A. S. Dzurak, and A. Morello, Nature **496**, 334 (2013).
- [28] G. Kirchmair, F. Zähringer, R. Gerritsma, M. Kleinmann, O. Gühne, A. Cabello, R. Blatt, and C. F. Roos, Nature **460**, 494 (2009).
- [29] A. Kubanek, M. Koch, C. Sames, A. Ourjoumtsev, P. W. H. Pinkse, K. Murr, and G. Rempe, Nature **462**, 898 (2009).
- [30] Th. Basché, S. Kummer, and C. Bräuchle, Nature **373**, 132 (1995).
- [31] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. O’Brien, A. G. White, and G. J. Pryde, PNAS **108**, 1256 (2011).
- [32] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. K. Shalm, and A. M. Steinberg, Science **332**, 1170 (2011).
- [33] A. Naik, O. Buu, M. D. LaHaye, A. D. Armour, A. A. Clerk, M. P. Blencowe, and K. C. Schwab, Nature **443**, 193 (2006).
- [34] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).

## Supplementary Information

Søren Gammelmark, Brian Julsgaard, and Klaus Mølmer  
*Department of Physics and Astronomy, Aarhus University,  
Ny Munkegade 120, DK-8000 Aarhus C, Denmark.*

(Dated: May 22, 2013)

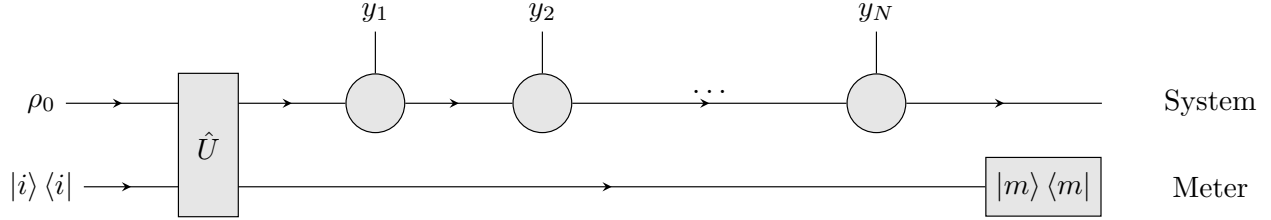


FIG. 1. Graphical illustration of the system dynamics considered in this note. Our open quantum system (top line) and a meter (bottom line) interact as described by the unitary operator  $\hat{U}$ . The system is probed by subsequent measurements with outcomes  $y_1, \dots, y_N$ , and later, the meter is subject to a projective measurement in some basis  $\{|m\rangle\}$ .

In this supplementary note we show that with our definition of the past quantum state we obtain identical results for past measurement outcomes as predicted by the ordinary quantum formalism, when this is applied to deferred measurements on a combined system-meter set-up [1]. Instead of an expression for determining past measurement outcomes propagating forward in time and including the meter system, we derive a backward propagating *effect matrix*  $E$  for the system only. We provide the specific stochastic differential equations for diffusion and jump type probing of the system (corresponding to homodyne detection and photon counting schemes). These equations can be chosen linear and non-trace preserving or non-linear and trace-preserving as the usual quantum filtering equations [2]. Finally, we demonstrate that our past quantum state generalizes the so-called *smoothed* state in classical hidden Markov models to the quantum case: The conditional density matrix  $\rho(t)$  is a natural generalization of the  $\alpha$ -state conditional probability, and the backward propagating effect matrix  $E$  is the quantum generalization of the backward propagating  $\beta$ -state in hidden Markov models [3].

## A. DEFERRED MEASUREMENTS AND THE PAST QUANTUM STATE

Imagine an observer using a meter to perform any measurement on our system which is initially prepared in a state which is represented by the density matrix  $\rho_0$ . The observer correlates the meter and our system by a unitary interaction,  $\hat{U}$ , between them. After this unitary interaction the system evolves independently of the meter and during this evolution, we measure our system a number of times  $N$  as illustrated in Fig. 1. By taking a suitable

limit with  $N \rightarrow \infty$  a description of continuous-time observation can be obtained. The observer who now has the meter in her possession can choose to perform a measurement of her choice at any later time. Our goal is to predict the result of such a measurement.

We have access to the results of the  $N$  measurements, where each measurement is in the most general case described by a set of measurement effect operators  $\hat{M}_y$  for each measurement outcome  $y$  in the set of possible measurement results  $Y$ . Using this formalism the effect of the measurement result  $y$  is to update a density matrix  $\rho$  according to  $\rho \xrightarrow{y} \hat{M}_y \rho \hat{M}_y^\dagger / \text{Tr}(\hat{M}_y^\dagger \hat{M}_y \rho)$ . The resulting density matrix, which we denote  $\rho|y$ , is conditioned upon the measurement result  $y$ . The probability of obtaining the result  $y$  is given by  $P(y) = \text{Tr}(\hat{M}_y^\dagger \hat{M}_y \rho)$ .

The operators  $\hat{M}_y$  should satisfy  $\sum_{y \in Y} \hat{M}_y^\dagger \hat{M}_y = \hat{I}$  such that the probability for observing any  $y \in Y$  is unity. If the observation is made with less than 100% readout efficiency the resulting conditioned density matrix can be written as a sum

$$\rho \xrightarrow{y} \frac{\sum_{k=1}^K \hat{M}_{k|y} \rho \hat{M}_{k|y}^\dagger}{\text{Tr} \left( \sum_{k=1}^K \hat{M}_{k|y}^\dagger \hat{M}_{k|y} \rho \right)}, \quad (\text{A1})$$

where  $\hat{M}_{k|y}$  are operators describing different possible effects which are all associated with the measurement result  $y$ . The number of terms,  $K$ , can depend on  $y$ .

Equation (A1) captures the effect of most types of dynamics of open quantum systems, including unitary evolution, dissipation effects described by a master equation on Lindblad form, and measurements. A density matrix subject to unitary evolution with unitary operator  $\hat{U}$  is thus updated according to  $\rho \mapsto \hat{U} \rho \hat{U}^\dagger$ . Where only the single unitary operator  $\hat{M}$  is needed clearly satisfies  $\hat{M}^\dagger \hat{M} = \hat{I}$ .

A projective measurement is described by a set of projectors  $\hat{\Pi}_a$  where  $a$  are eigenvalues of the observable  $\hat{A}$  being measured, and this case is also included by the identification  $\hat{M}_a = \hat{\Pi}_a$ . The effect on the system density matrix by a projective measurement is then simply the usual projection postulate  $\rho \xrightarrow{a} \hat{\Pi}_a \rho \hat{\Pi}_a / \text{Tr}(\hat{\Pi}_a \rho)$ . If we know that a projection measurement is performed, but the result of the measurement is hidden from us, the density matrix is updated according to  $\rho \mapsto \sum_a \hat{\Pi}_a \rho \hat{\Pi}_a$ . In this final example, all off-diagonal density matrix elements in the eigen basis of  $\hat{A}$  are zeroed, and the unobserved measurement operation is therefore equivalent to a decoherence process.

If a number of measurements are performed in sequence the density matrix  $\rho$  is updated



repeatedly by the formula Eq. (A1). In the limit of continuous-time measurements we can also describe the conditional time-evolution of the system density matrix. Consider for example a quantum system subject to homodyne detection. In this case the effect of a detection in a small interval of time  $dt$  is given by the operators

$$\hat{M}_{dY_t} = (2\pi dt)^{-1/4} \exp(-dY_t^2/4dt)(\hat{I} - i\hat{H}dt - \hat{c}^\dagger \hat{c}/2dt + \hat{c}dY_t), \quad (\text{A2})$$

where  $\hat{H}$  is the system Hamiltonian and  $\hat{c}$  is the system operator coupling to the real homodyne output signal  $dY_t$ . [2, 4] The probability for observing  $dY_t$  in this infinitesimal interval of time is a normal distribution with mean value  $\text{Tr}((\hat{c} + \hat{c}^\dagger)\rho)dt$  and variance  $dt$ . By applying the update formula Eq. (A1) with these operators we get Eq. (3) with  $\eta = 1$  and only one Lindblad operator  $\hat{L}_1 = \hat{c}$ . By including unobserved output channels for Lindblad operator  $\hat{L}_i$ ,  $i > 1$  and including limited detector efficiency  $\eta < 1$ , Eq. (3) turns out to be a special case of Eq. (A1).

Let us return to our main line of inquiry. The meter is assumed to be initialized in a pure state  $|i\rangle \langle i|$  when the observer applies the unitary interaction, acting on the combined system-meter state  $\rho_0 \otimes |i\rangle \langle i|$ . The resulting state is  $\hat{U}(\rho_0 \otimes |i\rangle \langle i|)\hat{U}^\dagger$ . By inserting complete bases of the meter  $\hat{I}_M = \sum_{m \in M} |m\rangle \langle m|$  we obtain

$$\rho = \sum_{m, m' \in M} \hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger \otimes |m\rangle \langle m'|, \quad (\text{A3})$$

where we have defined the system operators  $\hat{\Omega}_m = (\hat{I} \otimes \langle m|)\hat{U}(\hat{I} \otimes |i\rangle)$ . The operators  $\hat{\Omega}_m$ , defined this way, satisfy the requirement for measurement effect operators,  $\sum_{m \in M} \hat{\Omega}_m^\dagger \hat{\Omega}_m = \hat{I}$ , and with the chosen form for the operators  $\hat{\Omega}_m$ , we use the coupling to the meter to formally interrogate the properties of the open quantum system at the intermediate time  $t$ . We imagine that the open system dynamics proceeds, including the observations on the system continue with measurements result  $y \in Y$ , which cause the continued system dynamics described by Eq. (A1). In the general formula Eq. (A1) we can include unitary dynamics, dissipation channels and partially efficient measurements. As noted above, all these effects can be included by a suitable choice of the operators  $\hat{M}_{k|y}$ .

Since the subsequent dynamics only concern the system, the operators  $\hat{M}_{k|y}$  only act on the system degrees of freedom. The normalized system-meter state conditioned on one

measurement with the result  $y$  is

$$\rho|y = \frac{\sum_{k=1}^K \sum_{m,m' \in M} \hat{M}_{k|y} \hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger \hat{M}_{k|y}^\dagger \otimes |m\rangle \langle m'|}{\sum_{k=1}^K \sum_{m \in M} \text{Tr} \left( \hat{M}_{k|y} \hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger \hat{M}_{k|y}^\dagger \right)}. \quad (\text{A4})$$

We can calculate the expectation value of any meter-observable  $\hat{X}$  in the state  $\rho|y$  by the usual formalism,

$$\mathbb{E}[\hat{X}|y] = \text{Tr}(\hat{I} \otimes \hat{X} \rho|y) = \frac{\sum_{k=1}^K \sum_{m,m' \in M} \text{Tr}(\hat{M}_{k|y} \hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger \hat{M}_{k|y}^\dagger) \langle m'| \hat{X} |m\rangle}{\sum_{k=1}^K \sum_{m' \in M} \text{Tr} \left( \hat{M}_{k|y} \hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger \hat{M}_{k|y}^\dagger \right)}. \quad (\text{A5})$$

A projective measurement on the meter, which is now conditional on the system measurement result  $y$ , and yields the result  $m$  with the probability

$$P(m|y) = \frac{\sum_{k=1}^K \text{Tr}(\hat{M}_{k|y} \hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger \hat{M}_{k|y}^\dagger)}{\sum_{k=1}^K \sum_{m' \in M} \text{Tr} \left( \hat{M}_{k|y} \hat{\Omega}_{m'} \rho_0 \hat{\Omega}_{m'}^\dagger \hat{M}_{k|y}^\dagger \right)}, \quad (\text{A6})$$

where we have calculated the conditional expectation value of the meter projection operator  $\hat{X} = |m\rangle \langle m|$  to obtain the conditional probability  $P(m|y)$ . Note that we can rewrite the numerator in (A6) as

$$\sum_{k=1}^K \text{Tr}(\hat{M}_{k|y} \hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger \hat{M}_{k|y}^\dagger) = \text{Tr} \left( \hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger \left[ \sum_{k=1}^K \hat{M}_{k|y}^\dagger \hat{M}_{k|y} \right] \right) \equiv \text{Tr}(\hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger E), \quad (\text{A7})$$

where we have defined the *effect matrix*  $E = \sum_{k=1}^K \hat{M}_{k|y}^\dagger \hat{M}_{k|y}$ , and we obtain

$$P(m|y) = \frac{\text{Tr}(\hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger E)}{\text{Tr}(\sum_{m' \in M} \hat{\Omega}_{m'} \rho_0 \hat{\Omega}_{m'}^\dagger E)}. \quad (\text{A8})$$

This is the probability for the outcomes, stored "in the safe" as described in the main text. *I.e.*,  $P(m|y)$  is the retrodicted probability for the result  $m$  of a general measurement on the system with the corresponding measurement operator  $\{\hat{\Omega}_m\}$ . It notably differs from the usual formula  $P(m) = \text{Tr}(\hat{\Omega}_m^\dagger \hat{\Omega}_m \rho_0)$ , as it is possible to predict the outcome of measurements on the meter better than before we had access to the later measurement result  $y$ .

In the same way as the usual time dependent quantum state of a system, represented by a wave function  $\psi(t)$  or a density matrix  $\rho(t)$ , yields probabilities for general measurements on the system, we have now identified a mathematical structure, composed of  $\rho_0$  and  $E$ , which provide the probabilities for past measurements on a quantum system. We thus call the pair of matrices  $\Xi = (\rho_0, E)$  the *past quantum state*.

If the system has evolved until time  $t$ , and multiple measurements have been performed before the meter interacts with our system,  $\Xi(t) = (\rho(t), E(t))$ , where  $\rho(t)$  is the usual open system density matrix found by a stochastic equation of evolution conditioned on the measurements *before* time  $t$ . If multiple measurements are performed in sequence after the meter has interacted with our system at time  $t$ , the effect matrix  $E(t)$  depends on all measurement results *after* the coupling to the meter. In the following sections, we will derive efficient equations of evolution to determine  $E$  for both general probing scenarios and for a few special cases.

### 1. Dynamical equations for the effect matrix

Assume that the measurements up to time  $t$  has been taken into account in the forward state  $\rho_0$  then the generalization of Eq. (A5) when two subsequent measurements with result  $y_1$  and  $y_2$  are performed is

$$\mathbb{E}[\hat{X}|y_1, y_2] = \frac{\sum_{m, m' \in M} \sum_{k_1, k_2=1}^K \text{Tr} \left( \hat{M}_{k_2|y_2} \hat{M}_{k_1|y_1} \hat{\Omega}_m \rho_0 \hat{\Omega}_{m'}^\dagger \hat{M}_{k_1|y_1}^\dagger \hat{M}_{k_2|y_2}^\dagger \right) \langle m' | \hat{X} | m \rangle}{\sum_{m \in M} \sum_{k_1, k_2=1}^K \text{Tr} \left( \hat{M}_{k_2|y_2} \hat{M}_{k_1|y_1} \hat{\Omega}_m \rho_0 \hat{\Omega}_m^\dagger \hat{M}_{k_1|y_1}^\dagger \hat{M}_{k_2|y_2}^\dagger \right)}. \quad (\text{A9})$$

By using the cyclic property of the trace, the numerator can be written as

$$\sum_{m, m' \in M} \text{Tr} \left( \hat{\Omega}_m \rho_0 \hat{\Omega}_{m'}^\dagger \sum_{k_1=1}^K \left\{ \hat{M}_{k_1|y_1}^\dagger \left[ \sum_{k_2=1}^K \hat{M}_{k_2|y_2}^\dagger \hat{M}_{k_2|y_2} \right] \hat{M}_{k_1|y_1} \right\} \right) \langle m' | \hat{X} | m \rangle, \quad (\text{A10})$$

In this case, the effect matrix  $E$  is therefore given by

$$E = \sum_{k_1=1}^K \hat{M}_{k_1|y_1}^\dagger \left[ \sum_{k_2=1}^K \hat{M}_{k_2|y_2}^\dagger \hat{M}_{k_2|y_2} \right] \hat{M}_{k_1|y_1}, \quad (\text{A11})$$

where  $E$  now depends explicitly on the two future measurement results  $y_1$  and  $y_2$ .

From this we see that the update formula for  $E$ , as a counterpart to Eq. (A1), is given by the adjoint update

$$E \xrightarrow{y} \sum_{k=1}^K \hat{M}_{k|y}^\dagger E \hat{M}_{k|y} \quad (\text{A12})$$

where  $E$  equals the identity  $\hat{I}$  at the final time of measurements  $T$ , and is propagated recursively *backward* as indicated in the case of two measurements,

$$\hat{I} \xrightarrow{y_2} \sum_{k_2=1}^K \hat{M}_{k_2|y_2}^\dagger \hat{M}_{k_2|y_2} \xrightarrow{y_1} \sum_{k_1=1}^K \hat{M}_{k_1|y_1}^\dagger \left[ \sum_{k_2=1}^K \hat{M}_{k_2|y_2}^\dagger \hat{M}_{k_2|y_2} \right] \hat{M}_{k_1|y_1}. \quad (\text{A13})$$

The propagation is readily generalized to the case of  $N$  measurements,

$$E_N = \hat{I} \xrightarrow{y_N} E_{N-1} \xrightarrow{y_{N-1}} \dots \xrightarrow{y_1} E_0. \quad (\text{A14})$$

A hermitian operator  $E$  remains hermitian since the right hand side of Eq. (A12) is invariant under hermitian conjugation. Indeed,  $E$  has a separate physical interpretation as the positive semi-definite operator which given the state  $\rho_0$  yields the probability for the sequence of future measurement results,  $P(y_1, \dots, y_N | \rho_0) = \text{Tr}(E\rho_0)$ .

## 2. The past density matrix

Our definition of the past quantum state necessitates the use of two matrices from which probabilities can be generally determined via (A8). That expression may, however, be simplified in the special, but interesting, case where the system and meter are coupled very weakly.

We choose a two-dimensional quantum system as our meter, represented as a spin-1/2-particle, interacting briefly with our system by an interaction  $\hat{V} = ig(\hat{A}\hat{\sigma}^\dagger - \hat{A}^\dagger\hat{\sigma})$  where  $\hat{\sigma}$  is the spin lowering operator and  $\hat{A}$  is *any*, not necessarily hermitian, system operator. The meter is initially prepared in the spin-down state  $|\downarrow\rangle$ , and we allow the interaction to be active for a duration  $\tau$  such that  $\hat{U} = \exp(\epsilon(\hat{A}\hat{\sigma}^\dagger - \hat{A}^\dagger\hat{\sigma}))$  where  $\epsilon = \tau g$ . In the (weak) limit  $\epsilon \ll 1$

$$\hat{U} = \hat{I} + \epsilon(\hat{A}\hat{\sigma}^\dagger - \hat{A}^\dagger\hat{\sigma}) - \frac{\epsilon^2}{2} \left( \hat{A}\hat{A}^\dagger\hat{\sigma}^\dagger\hat{\sigma} + \hat{A}^\dagger\hat{A}\hat{\sigma}\hat{\sigma}^\dagger \right) + O(\epsilon^3). \quad (\text{A15})$$

If the meter is initialized in spin down in the  $z$ -direction  $|\downarrow\rangle$  then the measurement effect operators in the  $z$ -basis are given by  $\hat{\Omega}_\mu = (\hat{I} \otimes \langle\mu|)\hat{U}(\hat{I} \otimes |\downarrow\rangle)$ ,  $\mu = \downarrow, \uparrow$ ,

$$\hat{\Omega}_\downarrow = \hat{I} - \frac{\epsilon^2}{2}\hat{A}^\dagger\hat{A} + O(\epsilon^3) \quad (\text{A16})$$

$$\hat{\Omega}_\uparrow = \epsilon\hat{A} + O(\epsilon^3). \quad (\text{A17})$$

The effect of the measurement on the system state  $\rho_0$  when the result is not revealed is

$$\rho_0 \longmapsto \hat{\Omega}_\downarrow\rho_0\hat{\Omega}_\downarrow^\dagger + \hat{\Omega}_\uparrow\rho_0\hat{\Omega}_\uparrow^\dagger = \rho_0 + O(\epsilon^2). \quad (\text{A18})$$

The measurement associated with the subsequent readout of the meter is weak in the sense that it leaves the system undisturbed to first order in  $\epsilon$ . This type of measurement can

thus be performed unnoticed, and our formalism should enable prediction of its outcome in a seemingly "counter-factual" manner ("If  $\hat{A}$  had been measured, the outcome would have been ... "). The nature of the weak measurement allows an unknown agent to perform measurements without our knowledge and hide the result from us. At any later time, this person can inform us of the result, and we need the present theory to most accurately predict the outcome of this result. This ability comes with a cost, as to the interpretation of the weak measurement result, which we will return to shortly.

An arbitrary projective measurement of the meter is described by linear combinations of  $\hat{\Omega}_\downarrow$  and  $\hat{\Omega}_\uparrow$ . For measurements in the  $\hat{\sigma}_x$  and  $\hat{\sigma}_y$ -bases we can express the conditional expectation of  $\hat{\sigma}_x = (\hat{\sigma} + \hat{\sigma}^\dagger)$  and  $\hat{\sigma}_y = i(\hat{\sigma} - \hat{\sigma}^\dagger)$  by the real and imaginary parts of the expectation value of the step down operator  $\hat{\sigma}$ , respectively.

The expectation value of the meter operator  $\hat{\sigma}$  is

$$\mathbb{E}[\hat{\sigma}|y] = \frac{\sum_{k=1}^K \text{Tr}(\hat{M}_{k|y} \epsilon \hat{A} \rho_0 \hat{M}_{k|y}^\dagger)}{\sum_{k=1}^K \text{Tr}(\hat{M}_{k|y} \rho_0 \hat{M}_{k|y}^\dagger)}. \quad (\text{A19})$$

In this formula the denominator is independent of  $\hat{A}$ . This is due to the weak nature of the measurement, and it implies, that the expectation value is *linear* in the system operator  $\hat{A}$ . Since the meter couples to the system observable  $\hat{A}$ , it is natural to consider the weak value,

$$\langle A \rangle_w \equiv \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \mathbb{E}[\hat{\sigma}|y] = \frac{\sum_{k=1}^K \text{Tr}(\hat{M}_{k|y} \hat{A} \rho_0 \hat{M}_{k|y}^\dagger)}{\sum_{k=1}^K \text{Tr}(\hat{M}_{k|y} \rho_0 \hat{M}_{k|y}^\dagger)}. \quad (\text{A20})$$

We observe that to retrieve the weak value, i.e., the average outcome of a weak measurement, the information in the past quantum state  $\Xi = (\rho_0, E)$  with the effect matrix  $E = \sum_{k=1}^K \hat{M}_{k|y}^\dagger \hat{M}_{k|y}$  can be deduced from a *past density matrix*  $\rho_p$ . The expression  $\langle A \rangle_w = \text{Tr}(\hat{A} \rho_p)$  holds if we identify

$$\rho_p = \frac{\rho_0 E}{\text{Tr}(\rho_0 E)}. \quad (\text{A21})$$

It is worth noting that  $\rho_p$  is not Hermitian and even if  $\hat{A}$  is Hermitian,  $\text{Tr}(\hat{A} \rho_p)$  may have both real and imaginary parts. This is a well-known property of weak measurements, and it does not require the readout of, non-physical, complex measurement results, since the real part of  $\langle A \rangle_w$  refers to the (real) mean value of the  $\hat{\sigma}_x$ -operator of the meter, while its imaginary part is obtained by measuring the average  $\hat{\sigma}_y$ -spin component of the meter.

These averages may, in turn, indicate mean values of the system observable that are very different from its spectrum of eigenvalues. But this is well understood as an interference effect, when a measurement outcome is conditioned on different pre- and post-selected state of a physical system. [5]

### 3. Projective read-out measurements and the past quantum state

Imagine now that we perform a projective read-out measurement of the system observable  $\hat{A}$  using our meter. In this case  $\hat{\Omega}_m = \hat{\Pi}_{a_m}$  where the different measurement results are the eigenvalues of the observable  $\hat{A}$  denoted  $a_m$ . Following Eq. (A8) the probability for observing the eigenvalue  $a$  conditional on the later measurement result  $y$  is

$$P(a|y) = \frac{\text{Tr}(\hat{\Pi}_{a_m} \rho_{0-} \hat{\Pi}_{a_m} E_+)}{\text{Tr} \left( \sum_{m'} \hat{\Pi}_{a_{m'}} \rho_{0-} \hat{\Pi}_{a_{m'}} E_+ \right)}$$

where we by the  $+$  and  $-$  signs emphasize, that  $E_+$  is the effect matrix including measurements from immediately *after* the projective read-out measurement was performed until time  $T$  (here exemplified with a single measurement with result  $y$ ), and  $\rho_{0-}$  is the usual quantum state conditioned on the measurements until immediately *before* the projective read-out measurement was performed. The expectation value of the projective measurement of the observable  $\hat{A}$  is then

$$\mathbb{E}[\hat{A}|y] = \sum_m a_m P(a_m|y) = \frac{\text{Tr}(\sum_m a_m \hat{\Pi}_{a_m} \rho_{0-} \hat{\Pi}_{a_m} E_+)}{\text{Tr} \left( \sum_{m'} \hat{\Pi}_{a_{m'}} \rho_{0-} \hat{\Pi}_{a_{m'}} E_+ \right)}.$$

By inserting a resolution of the identity  $\hat{I} = \sum_{m'} \hat{\Pi}_{a_{m'}}$  this expression can be written in two ways

$$\mathbb{E}[\hat{A}|y] = \frac{\text{Tr} \left( \hat{A} \rho_{0-} \sum_m \hat{\Pi}_{a_m} E_+ \hat{\Pi}_{a_m} \right)}{\text{Tr} \left( \rho_{0-} \sum_m \hat{\Pi}_{a_m} E_+ \hat{\Pi}_{a_m} \right)} = \frac{\text{Tr} \left( \hat{A} \sum_m \hat{\Pi}_{a_m} \rho_{0-} \hat{\Pi}_{a_m} E_+ \right)}{\text{Tr} \left( \sum_m \hat{\Pi}_{a_m} \rho_{0-} \hat{\Pi}_{a_m} E_+ \right)}.$$

By introducing the effect matrix which takes the unobserved projective measurement into account by the map,  $E_- = \sum_m \hat{\Pi}_{a_m} E_+ \hat{\Pi}_{a_m}$  we can write this result

$$\mathbb{E}[\hat{A}|y] = \text{Tr}(\hat{A} \rho_{p-}), \quad (\text{A22})$$

defining the past density matrix as  $\rho_{p-} = \rho_{0-} E_- / \text{Tr}(\rho_{0-} E_-)$  immediately prior to the projective measurement.

Alternatively, we get

$$\mathbb{E}[\hat{A}|y] = \frac{\text{Tr}(\hat{A}\rho_{0+}E_+)}{\text{Tr}(\rho_{0+}E_+)}, \quad (\text{A23})$$

by using the past density matrix  $\rho_{p+} = \rho_{0+}E_+/\text{Tr}(\rho_{0+}E_+)$ , with the unobserved measurement process modifying the state  $\rho_{0+} = \sum_m \hat{\Pi}_{a_m}\rho_0\hat{\Pi}_{a_m}$ .

The projective measurement disturbs the system, and it must hence be taken into account in one of the components of the past quantum state: In Eq. (A22) it is included in the effect matrix  $E_-$  whereas in Eq. (A23) it is included in the forward density matrix  $\rho_{0+}$ . The two resulting alternative forms for the expectation value of  $\hat{A}$  are equivalent since  $\hat{A}$  commutes with the effect of the projective measurement.

While the general formula Eq. (A5) using the past quantum state  $\Xi = (\rho_0, E)$  can be applied to yield the probabilities for any past measurement outcome, it is interesting that the formalism may be brought closer to usual mean value expressions by use of the appropriately defined past density matrix.

#### 4. Differential equations for homodyne and counting measurements

A quantum system subject to homodyne or heterodyne detection satisfies an Itô stochastic differential equation of the form given in Eq. (3). As discussed previously such a measurement scenario fits into the present formulation by using the measurement operator for the time interval from  $t$  to  $t + dt$  is given by Eq. (A2) if we assume 100% efficiency and no unobserved channels, i.e. no dissipation effects. The adjoint update Eq. (A12) is then

$$dE_t \equiv E_{t-dt} - E_t = \left[ i \left[ \hat{H}, E_t \right] - \frac{1}{2} \left\{ \hat{c}^\dagger \hat{c}, E_t \right\} + \hat{c}^\dagger E_t \hat{c} \right] dt + \left[ \hat{c}^\dagger E_t + E_t \hat{c} \right] dY_{t-dt}. \quad (\text{A24})$$

where we have chosen a normalization of the effect matrix  $E$  such that the front factor  $(2\pi dt)^{-1/4} \exp(-dY^2/4dt)$  is discarded and  $E(T) = \hat{I}$  where  $T$  is the final time of observation.

By including dissipation effects and limited detector efficiency we obtain Eq. (4).

The conditional time evolution of a quantum system state subject to discrete counting signals in  $N$  channels can be described by the infinitesimal operators

$$\begin{aligned} \hat{M}_0 &= \hat{I} - i\hat{H}dt - \sum_{m=1}^N \frac{1}{2} \hat{L}_m^\dagger \hat{L}_m dt \\ \hat{M}_m &= \hat{L}_m \sqrt{dt} \quad \text{for } 1 \leq m \leq N \end{aligned} \quad (\text{A25})$$

where  $\hat{L}_m$  are Lindblad operators describing quantum jumps associated with the emission of quanta by the system into the environment.  $\hat{M}_0$  yields the measurement effect when no quantum (photon) is detected and  $\hat{M}_m$  indicates that a quantum is detected in environment channel number  $m$ .

If only the first channel is detected we get the well-established quantum jump filtering equation[2]

$$d\rho_t = \left[ -i [\hat{H}, \rho_t] + \sum_{m=2}^N \left( \hat{L}_m \rho_t \hat{L}_m^\dagger - \frac{1}{2} \{ \hat{L}_m^\dagger \hat{L}_m, \rho_t \} \right) - \frac{1}{2} \{ \hat{L}_1^\dagger \hat{L}_1, \rho_t \} + \text{Tr}(\hat{L}_1^\dagger \hat{L}_1 \rho_t) \rho_t \right] dt + \left[ \frac{\hat{L}_1 \rho_t \hat{L}_1^\dagger}{\text{Tr}(\hat{L}_1^\dagger \hat{L}_1 \rho_t)} - \rho_t \right] dN_t, \quad (\text{A26})$$

where  $dN_t = 0$  in all the intervals where no photon is detected, but  $dN_t = 1$  (and  $dt = 0$ ) at the instants of time a photon is detected in channel 1.

The adjoint update of the effect matrix is

$$dE_t \equiv E_{t+dt} - E_t = \left[ i [\hat{H}, E_t] + \sum_{m=2}^N \left( \hat{L}_m^\dagger E_t \hat{L}_m - \frac{1}{2} \{ \hat{L}_m^\dagger \hat{L}_m, E_t \} \right) - \frac{1}{2} \{ \hat{L}_1^\dagger \hat{L}_1, E_t \} \right] dt + \left[ \hat{L}_1^\dagger E_t \hat{L}_1 - E_t \right] dN_t. \quad (\text{A27})$$

## 5. Time evolution of mean values and weak values

For a quantum system subject to unit efficiency homodyne detection and no unobserved dissipation channels, the conventional mean value of a system observable  $\langle \hat{A} \rangle = \text{Tr}(\hat{A} \rho_t)$  changes with time according to the stochastic differential,

$$d\langle \hat{A} \rangle = -i \text{Tr}([\hat{A}, \hat{H}] \rho_t) dt + \text{Tr} \left( \hat{A} \hat{c} \rho_t \hat{c}^\dagger - \frac{1}{2} \{ \hat{A}, \hat{c}^\dagger \hat{c} \} \rho_t \right) dt + \left( \text{Tr}((\hat{A} \hat{c} + \hat{c}^\dagger \hat{A}) \rho_t) - \text{Tr}((\hat{c} + \hat{c}^\dagger) \rho_t) \text{Tr}(\hat{A} \rho_t) \right) dW_t, \quad (\text{A28})$$

where  $dW_t = dY_t - \text{Tr}((\hat{c} + \hat{c}^\dagger) \rho_t) dt$ .

We now address the similar change according to the past quantum state, i.e, the change of the weak value estimate  $\langle \hat{A} \rangle_w = \text{Tr}(\hat{A} \rho_p)$ . The differential of the un-normalized past density matrix is

$$d(\rho_t E_t) = \rho_{t+dt} E_{t+dt} - \rho_t E_t = d\rho_t E_{t+dt} - \rho_t dE_{t+dt}.$$



By inserting the expressions Eq. (A24) and Eq. (3) we get

$$d(\rho_t E_t) = -i [\hat{H}, \rho_t E_{t+dt}] dt + [\hat{c}, \rho_t \hat{c}^\dagger E_{t+dt}] dt - \frac{1}{2} [\hat{c}^\dagger \hat{c}, \rho_t E_{t+dt}] dt + [\hat{c}, \rho_t E_{t+dt}] dY_t. \quad (\text{A29})$$

Note that  $\text{Tr}(d(\rho_t E_t)) = 0$  as expected since  $\text{Tr}(\rho_t E_t) = \text{Tr}(\rho_T)$  is constant. The differential of the weak value  $\langle \hat{A} \rangle_w$ ,  $d\langle \hat{A} \rangle_w = d\text{Tr}(\hat{A} \rho_{p,t}) / \text{Tr}(\rho_t E_t) = \text{Tr}(\hat{A} d(\rho_t E_t)) / \text{Tr}(\rho_t E_t)$  thus becomes

$$d\langle \hat{A} \rangle_w = \frac{1}{\text{Tr}(\rho_t E_t)} \left[ -i \text{Tr}([\hat{A}, \hat{H}] \rho_t E_{t+dt}) dt + \text{Tr} \left( [\hat{A}, \hat{c}] \rho_t \hat{c}^\dagger E_{t+dt} - \frac{1}{2} [\hat{A}, \hat{c}^\dagger \hat{c}] \rho_t E_{t+dt} \right) dt + \text{Tr}([\hat{A}, \hat{c}] \rho_t E_{t+dt}) dY_t \right]. \quad (\text{A30})$$

This differs from the change in the conventional mean value and, in particular, we observe the suppression of the noise term  $\propto dY_t$  for observables which commute with the measurement operator  $\hat{c}$ . This suppression explains the smooth behavior of the black dashed curves in Figs. 2(a) and 2(b), while the conventional mean values  $\langle \hat{A} \rangle$  show fluctuations due to the Wiener noise term,  $dW$ , in Eq. (A28).

## B. RELATION TO CLASSICAL HIDDEN MARKOV MODELS

The classical theory of hidden Markov models is very well established and an excellent introduction can be found in Ref. 3. Here, a *hidden Markov model* is a discrete-time stochastic process where a hidden system state evolves according to a Markov chain and observations of some output signal depending on the system state are performed. Let  $X_t$  be the system state at time  $t$  and the output signal at time  $t$  be  $Y_t$ . The output signal  $Y_t$  depends only on the system state at the same time and its probability distribution is therefore completely determined by  $P(Y_t|X_t)$ . Since the evolution of  $X_t$  follows a Markov chain the system dynamics is completely determined by the transition probabilities  $P(X_{t+1}|X_t)$ . It is a simple matter to generalize these ideas to the continuous time case (such as a system governed by a rate equation), but for simplicity we will consider only the discrete time case here and the times  $t$  are integers.

The full joint probability distribution for measurements  $Y_t$  and states for a process running from time  $t = 0$  to time  $t = T$  is then

$$P(X_0, \dots, X_T, Y_1, \dots, Y_T) = \prod_{t=0}^{T-1} P(X_{t+1}|X_t) \prod_{t=1}^T P(Y_t|X_t) P(X_0), \quad (\text{B31})$$

where  $P(X_0)$  is the probability for the initial state  $X_0$ . The theory of hidden Markov models provides numerically efficient algorithms for calculating the following quantities: (1) the forward filtered state probability distribution  $P(X_t|Y_1, \dots Y_t)$  and (2) the smoothed state probability distribution  $P(X_t|Y_1, \dots Y_T)$  for all times  $t$ .

The forward estimate is easily calculated by a standard recursive Bayesian procedure. Following the notation in [3] we define the vectors

$$\alpha_t(i) = P(Y_1, \dots Y_t, X_t = i) \quad (\text{B32})$$

$$\beta_t(i) = P(Y_{t+1}, \dots Y_N | X_t = i). \quad (\text{B33})$$

Using the  $\alpha_t$  and  $\beta_t$ -vectors we can calculate the filtered and smoothed distributions at time  $t$  by the following formulas

$$P(X_t = i | Y_1, \dots Y_t) = \frac{\alpha_t(i)}{\sum_k \alpha_t(k)}, \quad (\text{B34})$$

$$P(X_t = i | Y_1, \dots Y_N) = \frac{\alpha_t(i)\beta_t(i)}{\sum_k \alpha_t(k)\beta_t(k)}, \quad (\text{B35})$$

both of which are variations of Bayes' formula. It is not difficult to show that  $\alpha_t$  and  $\beta_t$  satisfy the following recursion relations

$$\alpha_{t+1}(i) = \sum_j P(Y_{t+1} | X_{t+1} = i) P(X_{t+1} = i | X_t = j) \alpha_t(j) \quad (\text{B36})$$

$$\beta_t(i) = \sum_j P(Y_{t+1} | X_{t+1} = j) P(X_{t+1} = j | X_t = i) \beta_{t+1}(j), \quad (\text{B37})$$

where  $\beta_T(i) = 1$  and  $\alpha_0(i) = P(X_0 = i)$ . In the following, we show that our effect matrix is equivalent to the  $\beta_t$ -vector and that the past quantum state  $\Xi(t)$  is the natural quantum generalization of the hidden Markov model pair  $(\alpha_t, \beta_t)$ . In the case of non-disturbing measurements the hidden Markov model smoothed state (B35) is equivalent to both the past quantum state  $\Xi$  and the past density matrix  $\rho_p$ . This is consistent with the weak value assumption of no disturbance of the system due to the measurements since classical measurements may always be thought of as non-disturbing).

Indeed, the above hidden Markov theory can be formulated using diagonal density matrices and updates of the form given in Eq. (A1). Let  $|i\rangle$  be an orthogonal basis for a Hilbert space, where  $i$  denotes the same internal states as in the hidden Markov model. The Markov chain evolution  $P(X_{t+1}|X_t)$  is now given by the update

$$\mathcal{C}: \rho \longmapsto \sum_{i,j} |j\rangle \langle j| P(X_{t+1} = j | X_t = i) \langle i | \rho | i \rangle, \quad (\text{B38})$$

which is in fact an evolution of the type given in Eq. (A1) with  $\hat{M}_{i,j} = \sqrt{P(X_{t+1}=j|X_t=i)} |j\rangle \langle i|$ . The observation process is given by the update

$$\mathcal{I}: \rho \xrightarrow{y} \sum_i P(Y_t = y|X_t = i) \langle i|\rho|i\rangle |i\rangle \langle i| \quad (\text{B39})$$

followed by re-normalization as in Eq. (A1).

The classical hidden Markov model is then reproduced by picking the initial state  $\rho_0 = \sum_i P(X_0 = i) |i\rangle \langle i|$  and applying the two updates  $\mathcal{C}$  and  $\mathcal{I}$  in sequence.

The (un-normalized) forward filtered state  $\alpha_t$  is simply the (un-normalized) filtered quantum state  $\tilde{\rho}_t$  which satisfies the recursion relation

$$\tilde{\rho}_t \xrightarrow{\mathcal{C}} \sum_{i,j} |i\rangle \langle i| P(X_{t+1} = i|X_t = j) \langle j|\tilde{\rho}_t|j\rangle \quad (\text{B40})$$

$$\xrightarrow{\mathcal{I}} \sum_{i,j} |i\rangle \langle i| P(Y_{t+1} = y|X_{t+1} = i) P(X_{t+1} = i|X_t = j) \langle j|\tilde{\rho}_t|j\rangle \equiv \tilde{\rho}_{t+1} \quad (\text{B41})$$

which for diagonal  $\tilde{\rho}_t$  is exactly Eq. (B36).

The effect matrix  $E_t$  is initially  $E_T = \hat{I}$  which is equivalent to the initial condition  $\beta_T(i) = 1$ . The effect matrix is propagated according to the adjoint update Eq. (A12) as

$$E_{t+1} \xrightarrow{\mathcal{I}^\dagger} \sum_j |j\rangle \langle j| P(Y_{t+1} = y|X_{t+1} = j) \langle j|E_{t+1}|j\rangle \quad (\text{B42})$$

$$\xrightarrow{\mathcal{C}^\dagger} \sum_{i,j} |i\rangle \langle i| P(Y_{t+1} = y|X_{t+1} = j) P(X_{t+1} = j|X_t = i) \langle j|E_{t+1}|j\rangle \equiv E_t. \quad (\text{B43})$$

which is exactly the update formula for  $\beta_t$  Eq. (B37). Note here that the  $\mathcal{I}$ -update is unchanged, whereas the adjoint update for  $\mathcal{C}$  is modified.

The theory of hidden Markov models includes numerically efficient algorithms for re-estimating the parameters occurring in the model. The so-called Baum-Welch algorithm is a special case of the Expectation-Maximization algorithm. By using the fully smoothed estimate as encoded in the  $\alpha$ - and  $\beta$ -vectors, a simple formula for the re-estimated parameters exists, which leads to a more likely sequence of measurement results  $Y_1, \dots, Y_T$  given the model. In this way a local maximum of the likelihood can be calculated by iterating parameter re-estimation and the smoothing calculation. We believe that a similar technique can be applied to estimate unknown parameters in quantum processes using the past quantum

state.

- 
- [1] R. B. Griffiths and C.-S. Niu, *Phys. Rev. Lett.* **76**, 3228 (1996).
  - [2] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, New York, 2010).
  - [3] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (Cambridge University Press, New York, 2007).
  - [4] K. Jacobs and D. A. Steck, *Contemp. Phys.* **47**, 279 (2006).
  - [5] Y. Aharonov, D. Z. Albert, and L. Vaidman, *Phys. Rev. Lett.* **60**, 1351 (1988).